UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P O Box 1450
Alexandria, Virginia 22313-1450
www uspto gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/649,382	08/27/2003	Nebojsa Jojic	301911.01	6064

27662 7590 12/04/2008
MICROSOFT CORPORATION
C/O LYON & HARR, LLP
300 ESPLANADE DRIVE
SUITE 800
OXNARD, CA 93036

EXAMINER	
RASHID, DAVID	

ART UNIT	PAPER NUMBER
2624	

MAIL DATE	DELIVERY MODE
12/04/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	**Application No.**	**Applicant(s)**	
	10/649,382	JOJIC ET AL.	
	Examiner	**Art Unit**	
	DAVID P. RASHID	2624	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *22 September 2008*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-32* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-32* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Table of Contents

Continued Examination Under 37 CFR 1.114

[1] A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on September 22, 2008 has been entered.

Amendments

[2] This office action is responsive to <u>After Final Amendment</u> received on August 20, 2008. Claims 1-32 remain pending.

Response to Arguments

[3] Remarks filed August 20, 2008 with respect to claims 1-32 have been respectfully and fully considered, but not found persuasive.

The Rejections of Claims Under 35 U.S.C. 102(b)

Summary of Remarks

> Foote does not teach the applicants' claimed preferred number of classes of objects to be identified within the image sequence or automatically decomposing the image sequence into the preferred number of classes of objects in near real-time. Nor does Foote teach in near-real time automatically decomposing each image sequence into a generative model including a set of model parameters comprising the mean visual appearance and variance of each class in the image sequence.
>
> . . .

Applicant's Remarks at 12-13, August 20, 2008.

Examiner's Response

However, "in near real-time" is highly subjective as there is nothing definite in the claim as to point about what degree constitutes being "near real-time" (whether it is less than a second, a matter of multiple seconds, minutes, hours, days etc). Foote only "segment[ing] a video file with corresponding audio after it has been recorded, not in real-time as it is being input" (Remarks at 13) is an extended limitation to what characteristics of Applicant's invention define more of "near real-time" that is not positively recited in the claim. The Examiner suggests amending the claim such that "segment[ing] a video file with corresponding audio after it has been recorded" cannot be read into the claim. As the claim stands now, segmenting video and audio after recording is all considered "in near real-time", whether it is a matter of milliseconds, seconds, hours, etc.

> It merely appears to determine video features in image frames and using these features to determine which of the predefined classes a frame belongs to. It does not teach automatically decomposing each image sequence into a generative model (e.g., a model of how the observed data could have been generated) with each generative model including a set of model parameters that represent at least one object class for each image sequence using an expectation-maximization analysis that employs a Viterbi analysis

Remarks at 13.

Examiner's Response

However, *Foote et al.* does disclose automatically decomposing (fig. 2, item 208; fig. 12, items 1202-1203; the final outcome at fig. 25 of three classes G,A,B) the image sequence (fig. 2, item 201) into the preferred number of classes of objects in near real-time ("segmenting. . .into a pre-defined set of classes" in 5:14-16; the "[t]raining [d]ata" and "[t]est [d]ata" of TABLE 1 at col. 12). "[S]egmenting. . .into a pre-defined set of classes" (*Foote et al.* at 5:14-16) is equivalent to "decomposing. . .into the preferred number of classes". TABLE 1 at Col. 12 of *Foote et al.* is a depiction of the predefined classes (of which were preferred) that include slides, crowd, longsw, longsb, etc. The training data and test date are both put into these classes (and additionally all the data is if you add the number of data in each class to equal to the total amount). Fig. 25 is another depiction that all data ends up in either class G, A, or B. This number of classes was also "preferred" if the algorithm was written to incorporate these three classes (as opposed to an algorithm that randomly selects the number of classes, and thus not preferred). In addition, G, A, and B are collectively a definite "number".

The Rejections of Claims Under 35 U.S.C. 103(a)

Summary of Remarks

> As discussed above Foote does not teach the applicants' claimed preferred number of classes of objects to be identified within the image sequence or automatically decomposing the image sequence into the preferred number of classes of objects in near real-time. Nor does Foote teach in near-real time automatically decomposing each image sequence into a generative model including

> a set of model parameters comprising the mean visual appearance and variance of each class in the image sequence. Petrovic also does not teach these features.
>
> . . .

Remarks at 16.

"Dellaert also does not teach these features." Remarks at 18. "Dellaert and Eberman also do not teach these features." Remarks at 19. "Jojic also does not teach these features." Remarks at 21. "Eberman also does not teach these features." Remarks at 22.

Examiner's Response

However, as shown above *Foote et al.* does disclose a "preferred number of classes of objects to be identified within the image sequence or automatically decomposing the image sequence into the preferred number of classes of objects in near real-time". *Petrovic, Dellart, Jojic*, and *Eberman* all do not need to teach these features.

Claim Rejections - 35 USC § 112

[4] The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

Claims 1-32 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Lack of Antecedent Basis

Claims 1-32 recite the limitation "the mean visual appearance and variance of each class" in claim 1, l. 10-11 and claim 23, l. 9-10. There is insufficient antecedent basis for this limitation in the claim. Claims 2-22 and 24-32 are rejected for failing to alleviate their dependent's deficiency.

Claim Rejections - 35 USC § 101

[5] 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

In Re Bilski – "Tied To" Criteria

With respect to **claims 1-32,** while the claims recite a series of steps or acts to be performed, a statutory "process" under 35 U.S.C. 101 must (1) be tied to another statutory category (such as a particular apparatus), or (2) transform underlying subject matter (such as an article or material) to a different state or thing. *See* Clarification of "Processes" under 35 U.S.C. 101, Deputy Commissioner for Patent Examining Policy, John J. Love, May 15, 2008; *available at* http://www.uspto.gov/web/offices/pac/dapp/opla/preognotice/section_101_05_15_2008.pdf.

The instant claims neither transform underlying subject matter nor positively tie to another statutory category that accomplishes the claimed method steps, and therefore do not qualify as a statutory process.

Claim Rejections - 35 USC § 102

[6] The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the

> international application designated the United States and was published under Article 21(2) of such treaty in the English language.

Foote et al.

[7] **Claims 1-3, 5-6, 14, 18-19,** and **23-24** are rejected under 35 U.S.C. 102(b) as being anticipated by U.S. Patent No. 6,404,925 (issued Jun. 11, 2002, hereinafter "Foote et al.").

Regarding **claim 1,** *Foote et al.* discloses a system (fig. 1; fig. 2) for automatically decomposing an image sequence (fig. 2, item 201), comprising a computer-readable storage medium (fig. 1, items 103, 107-108) storing a program that when executed (fig. 1, items 102, 109) performs the following process actions:

providing an image sequence (fig. 2, item 201) of at least one image frame (fig. 3, items 301-308) of a scene (e.g., scene in fig. 3);

providing only a preferred number of classes of objects (fig. 2, items 202-205; "pre-defined set of classes" in 5:14-16; "[e]xamples of video classes include close-ups of people, crowd scenes, and shots of presentation material. . ." at 5:16-20; "[s]hot [c]ategory" of TABLE 1 at col. 12) to be identified (fig. 12, item 1204) within the image sequence;

automatically decomposing (fig. 2, item 208; fig. 12, items 1202-1203; the final outcome at fig. 25 of three classes G,A,B) the image sequence (fig. 2, item 201) into the preferred number of classes of objects in near real-time ("segmenting. . .into a pre-defined set of classes" in 5:14-16; the "[t]raining [d]ata" and "[t]est [d]ata" of TABLE 1 at col. 12),

using probabilistic inference (fig. 23; "<u>hidden Markov model</u> to be used in the method for classifying a video according to the present invention. Each of the image classes G, A, and B, are modeled using Gaussian distributions" (emphasis added) at 16:49-55; computing posterior distribution of variables using the hidden Markov models; "[t]he similarity between a given

frame and the query is computed during the Viterbi algorithm as the posterior probability of the query state or states" at 18:42-44) and learning ("learning of the actual data points" at 15:34) to compute a single set of model parameters (the single set of mean visual appearances and variances in the "Gaussian distributions" at 6:32-33) comprising the mean visual appearance and variance (e.g., "Gaussian distributions having different means and variances" at 7:59-60; fig. 4; i.e., the model parameters from the hidden Markov model comprise means and variances of each class) of each class (fig. 2, items 202-205; "pre-defined set of classes" in 5:14-16; "[e]xamples of video classes include close-ups of people, crowd scenes, and shots of presentation material. . ." at 5:16-20; "[s]hot [c]ategory" of TABLE 1 at col. 12) in the image sequence (fig. 2, item 201).

In summary, the hidden Markov model of fig. 23 (probabilistic inference and learning) uses/computes Gaussian distributions (that compute a single set of model parameters comprising mean visual appearance and variance of each class in the image sequence). Doing this automatically decomposes the image sequence into the preferred number of classes of objects in near real time as shown in fig. 23, 25 (classes A,B,G).

Regarding **claim 2**, *Foote et al.* discloses the system of claim 1 wherein providing the preferred number of objects ("pre-defined set of classes" in 5:14-16) comprises specifying the preferred number of classes of objects via a user interface (a user interface is visual interface from which a user can interact with such as fig. 22; a pre-defined set of classes suggests that some sort of user interface must have been used to "define" the set of classes; "[t]he feature used for classification are general, so that users can define arbitrary class types" in 5:18-20).

Regarding **claim 3**, *Foote et al.* discloses the system of claim 1 wherein decomposing the image sequence (fig. 2, item 201) into the preferred number of objects ("segmenting...into a pre-

defined set of classes" in 5:14-16) comprises automatically learning a 2-dimensional model (fig. 3, items 310-322) of each object class (7:13-15).

Regarding **claim 5**, *Foote et al.* discloses the system of claim 1 wherein automatically decomposing the image sequence (fig. 2, item 201) into the preferred number of object classes ("pre-defined set of classes" in 5:14-16) comprises performing an inferential probabilistic analysis (fig. 2, items 202-205; "Gaussian distributions" in 5, line 65-6, line 2) of each image frame for identifying ("segmenting…into a pre-defined set of classes" in 5:14-16) the preferred number of object class appearances within the image sequence.

Regarding **claim 6**, *Foote et al.* discloses the system of claim 5 wherein performing an inferential probabilistic analysis of each image frame comprises performing a variational generalized expectation-maximization analysis (21:55-62) of each image frame (fig. 3, items 301-308) of the image sequence (fig. 2, item 201), wherein the expectation-maximization analysis employs a Viterbi algorithm (6:43-45; 16:40-42) in a process of filling in values of hidden variables (21:55-62; variables in fig. 4) in a model describing the object class.

Regarding **claim 14**, *Foote et al.* discloses the system of claim 1 wherein automatically decomposing the image sequence into the preferred number of object classes comprises performing a probabilistic variational expectation-maximization analysis (21:55-62).

Regarding **claim 18**, *Foote et al.* discloses the system of claim 1 further comprising a generative model ("hidden Markov model" in 18:35-42) which includes a set of model parameters ("alignment" in 18:35-42) that represent the entire image sequence ("entire video" in 18, line 37).

Regarding **claim 19**, *Foote et al.* discloses the system of claim 1 further comprising a generative model which includes a set of model parameters that represent the images of the image sequence processed to that point (21:4-15).

Regarding **claim 22**, *Foote et al.* discloses the system of claim 19 further comprising automatically reconstructing a representation of the image sequence from the generative model, wherein the representation comprises the preferred number of object classes (fig. 2, item 207).

Regarding **claim 23**, *Foote et al.* discloses a computer-implemented process (fig. 1; fig. 2) for automatically generating a representation of an object (e.g., "crowed" at TABLE 1, Col. 12) in at least one image sequence (fig. 2, item 201), comprising using a computer (fig. 1, items 103, 107-108) to:

acquire at least one image sequence (fig. 2, item 201), each image sequence having at least one image frame (fig. 3, items 301-308);

in near real-time automatically decompose each image sequence (fig. 2, item 201) into a generative model (fig. 2, items 202-205; "Gaussian distributions" in 5, line 65-6, line 2), with each generative model comprising a set of model parameters (the single set of mean visual appearances and variances in the "Gaussian distributions" at 6:32-33) comprising the mean visual appearance and variance (e.g., "Gaussian distributions having different means and variances" at 7:59-60; fig. 4; i.e., the model parameters from the hidden Markov model comprise means and variances of each class) of each class (fig. 2, items 202-205; "pre-defined set of classes" in 5:14-16; "[e]xamples of video classes include close-ups of people, crowd scenes, and shots of presentation material. . ." at 5:16-20; "[s]hot [c]ategory" of TABLE 1 at col. 12) in the image sequence (fig. 2, item 201) being decomposed using an expectation-maximization analysis

(fig. 23; "hidden Markov model to be used in the method for classifying a video according to the present invention. Each of the image classes G, A, and B, are modeled using Gaussian distributions" (emphasis added) at 16:49-55; computing posterior distribution of variables using the hidden Markov models; "[t]he similarity between a given frame and the query is computed during the Viterbi algorithm as the posterior probability of the query state or states" at 18:42-44) that employs a Viterbi analysis (6:43-45; 16:40-42).

Regarding **claim 24**, claim 2 recites identical features as in claim 24. Thus, references/arguments equivalent to those presented above for claim 2 are equally applicable to claim 24.

Claim Rejections - 35 USC § 103

[10] The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Foote et al. in view of Petrovic et al.

[11] **Claims 4, 7,** and **27** are rejected under 35 U.S.C. 103(a) as being unpatentable over *Foote et al.* in view of Transformed Hidden Markov Models: Estimating Mixture Models of Images and Inferring Spatial Transformations in Video Sequences, Computer Visions and Pattern Recognition, 2000, Vol. 2, pg 26 – 33 (hereinafter "Petrovic et. al").

Regarding **claim 4**, while *Foote et al.* discloses the system of claim 3, *Foote et al.* does not directly suggest wherein the model employs a latent image and a translation variable in learning each object class.

Petrovic et al. discloses transformed hidden markov model wherein the model employs a latent image ("latent image", pg 27-28) and a translation variable ("set of transformations…", pg 27, right column) in learning each object class.

It would have been obvious to one of ordinary skill in the art at the time the invention was made for the model of *Foote et al.* to employ a latent image and a translation variable in learning each object class as taught by *Petrovic et al.* to "develop a general video analysis tool that extracts long and short term similarities in video using a novel generative model, called the transformed hidden Markov model (THMM).", *Petrovic et al.*, pg 26 and to "learn models of different types of object from unlabeled frames in a video sequence that include background clutter, occlusion and spatial transformations, such as translation, rotation and shearing.", *Petrovic et al.*, pg. 26.

Regarding **claim 5**, while *Foote et al.* discloses the system of claim 3, *Foote et al.* does not directly suggest wherein the model describing the object class employs a latent image and a translation variable in filling in said hidden variables.

Petrovic et al. discloses transformed hidden markov model wherein the model describing the object class employs a latent image ("latent image", pg 27-28) and a translation variable ("set of transformations…", pg 27, right column) in filling in hidden variables (pg 29).

It would have been obvious to one of ordinary skill in the art at the time the invention was made for the model of *Foote et al.* to employ a latent image and a translation variable in filling in hidden variables as taught by *Petrovic et al.* to "develop a general video analysis tool that extracts long and short term similarities in video using a novel generative model, called the transformed hidden Markov model (THMM).", *Petrovic et al.*, pg 26 and to "learn models of

different types of object from unlabeled frames in a video sequence that include background clutter, occlusion and spatial transformations, such as translation, rotation and shearing.", *Petrovic et al.*, pg. 26.

Regarding **claim 27**, claim 4 recites identical features as in claim 27. Thus, references/arguments equivalent to those presented above for claim 4 are equally applicable to claim 27.

Foote et al. in view of Dellaert

[12] **Claims 8-10, 13, 15-17,** and **28-31** are rejected under 35 U.S.C. 103(a) as being unpatentable over *Foote et al.* in view of The Expectation Maximization Algorithm, College of Computing, Georgia Institute of Technology, Technical Report number GIT-GVU-02-20, 2/2002 (hereinafter "Dellaert").

Regarding **claim 8**, while *Foote et al.* discloses a generalized expectation-maximization analysis, *Foote et al.* does not directly teach wherein an expectation step of the generalized expectation-maximization analysis maximizes a lower bound on a log-likelihood of each image frame by inferring approximations of variational parameters.

Dellaert discloses the expectation maximization algorithm that teaches wherein an expectation step of the generalized expectation-maximization analysis maximizes a lower bound on a log-likelihood by inferring approximations of variational parameters (Section 2, "EM as Lower Bound Maximization").

It would have been obvious to one of ordinary skill in the art at the time the invention was made for the generalized expectation-maximization for each image frame of *Foote et al.* to include wherein an expectation step of the generalized expectation-maximization analysis

maximizes a lower bound on a log-likelihood by inferring approximations of variational parameters as taught by *Dellaert* as "[t]he goal is to maximize the posterior probability (1) of the parameters Θ given the data U, in the presence of hidden data J.", *Dellaert*, Section 2, "EM as Lower Bound Maximization".

Regarding **claim 9**, while *Foote et al.* discloses a generalized expectation-maximization analysis, *Foote et al.* does not directly teach wherein the maximization step of the generalized expectation-maximization analysis automatically adjusts model parameters in order to maximize a lower bound on a log-likelihood of each image frame.

Dellaert discloses the expectation maximization algorithm that teaches wherein the maximization step of the generalized expectation-maximization analysis automatically adjusts model parameters in order to maximize a lower bound on a log-likelihood (converting Θ into Θ^{t+1} in equation (4) in Section 2.2, "Maximizing the Bound").

It would have been obvious to one of ordinary skill in the art at the time the invention was made for the generalized expectation-maximization for each image frame of *Foote et al.* to include wherein the maximization step of the generalized expectation-maximization analysis automatically adjusts model parameters in order to maximize a lower bound on a log-likelihood as taught by *Dellaert* as "[t]he goal is to maximize the posterior probability (1) of the parameters Θ given the data U, in the presence of hidden data J.", *Dellaert*, Section 2, "EM as Lower Bound Maximization".

Regarding **claim 10**, while *Foote et al.* discloses a generalized expectation-maximization analysis, *Foote et al.* does not teach wherein the expectation step and the maximization step are performed once for each image in said image sequence.

Dellaert discloses the expectation maximization algorithm that teaches wherein the expectation step and the maximization step are performed once for each set of new data (equation (4) pg 6 to obtain Θ^{t+1} is only computed once for each set of new data).

It would have been obvious to one of ordinary skill in the art at the time the invention was made for each image frame of the image sequence of *Foote et al.* to be the new data as taught by Dellaert as "[t]he goal is to maximize the posterior probability (1) of the parameters Θ given the data U, in the presence of hidden data J.", *Dellaert*, Section 2, "EM as Lower Bound Maximization".

Regarding **claim 13**, *Foote et al.* discloses wherein automatic computation of the expectation step is accelerated by using a Viterbi analysis (6:43-45; 16:40-42; 18:31-48).

Regarding **claim 15**, while *Foote et al.* discloses a generalized expectation-maximization analysis, *Foote et al.* does not directly teach wherein the expectation-maximization analysis comprises: forming a probabilistic model having variational parameters representing posterior distributions; initializing said probabilistic model; inputting an image frame from the image sequence; computing a posterior given observed data in said image sequence; and using the posterior of the observed data to update the probabilistic model parameters.

Dellaert discloses the expectation maximization algorithm that teaches wherein the expectation-maximization analysis comprises:

forming a probabilistic model having variational parameters ("Θ^t", "Θ^{t+1}", means "θ_1" and "θ_2") representing posterior distributions (last paragraph, pg 1);

initializing said probabilistic model (the probabilistic model has to be initialized at some point to obtain Θ^{t+1});

inputting new data ("current guess" Θ^t from equation (3), pg 5 to "improved estimate" Θ^{t+1});

computing a posterior given observed data ("log-posterior log $P(\Theta|U)$", pg 6); and

using the posterior of the observed data to update the probabilistic model parameters ("M-step" equation, pg 6).

It would have been obvious to one of ordinary skill in the art at the time the invention was made for the new image frame from the image sequence of *Foote et al.* to be the new data as taught by *Dellaert* and that the generalized expectation-maximization analysis of *Foote et al.* to include wherein the expectation-maximization analysis comprises: forming a probabilistic model having variational parameters representing posterior distributions; initializing said probabilistic model; inputting; computing a posterior given observed data; and using the posterior of the observed data to update the probabilistic model parameters as taught by *Dellaert* as "[t]he goal is to maximize the posterior probability (1) of the parameters Θ given the data U, in the presence of hidden data J.", *Dellaert*, Section 2, "EM as Lower Bound Maximization".

Regarding **claim 16,** *Foote et al.* discloses wherein the expectation-maximization analysis further comprises:

outputting the model parameters (21:55-62).

Regarding **claim 17,** *Foote et al.* discloses further comprising incrementing to the next image frame in said image sequence and repeating the actions after initializing the probability model until the end of the image sequence has been reached (the loops in fig. 12, fig. 20, fig. 26, and fig. 28 until frame sequence are complete).

Regarding **claim 28**, claim 8 recites identical features as in claim 28. Thus, references/arguments equivalent to those presented above for claim 8 are equally applicable to claim 28.

Regarding **claim 29**, claim 9 recites identical features as in claim 29. Thus, references/arguments equivalent to those presented above for claim 9 are equally applicable to claim 29.

Regarding **claim 30**, claim 15 recites identical features as in claim 30. Thus, references/arguments equivalent to those presented above for claim 15 are equally applicable to claim 30.

Regarding **claim 31**, claim 16 recites identical features as in claim 31. Thus, references/arguments equivalent to those presented above for claim 16 are equally applicable to claim 31.

Foote et al. in view of Dellaert and Eberman et al.

[13] Claims 11-12 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Foote et al.* in view of *Dellaert* and U.S. Patent No. 5,925,065 (issued Jul. 13, 1999, hereinafter "Eberman et al.").

Regarding **claims 11** and **12**, while *Foote et al.* in view of *Dellaert* disclose a computer-readable process of claim 8 wherein computation of the expectation step is suggested to use some form of transform, *Foote et al.* in view of *Dellaert* does not teach accelerating the expectation step using a FFT-based inference analysis.

Eberman et al. teaches using a FFT-based inference analysis (5:19-27).

It would have been obvious for the computation of the expectation step of *Foote et al.* in view of *Dellaert* to include using a FFT-based inference analysis as taught by *Eberman et al.* to reduce calculation time ($2N^2$) as less computation is needed ($2 N \log_2 N$) as well known to one of ordinary skill in the art.

It is well known to one of ordinary skill in the art that using the FFT requires performance on variables (x_n, k, N) that are converted into a coordinate system (X_k coordinate system) wherein transforms applied to those variables are represented by shift operations (x_n shifted by exponential on right side of equation to equal X_k).

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} nk} \qquad k = 0, \ldots, N-1.$$

<u>*Foote et al. in view of Jojic et al.*</u>

[14] **Claims 20-21** and **25-26** are rejected under 35 U.S.C. 103(a) as being unpatentable over *Foote et al.* in view of <u>Learning Flexible Sprites in Video Layers</u>, Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, 2001, pg 1-8 (hereinafter "Jojic et al.").

Regarding **claim 20**, while *Foote et al.* discloses the system of 19, *Foote et al.* does not teach wherein the model parameters include: a prior probability of at least one object class; and means and variances of object appearance maps.

Jojic et al. teaches a learning flexible sprites in video layers wherein the model parameters include:

a prior probability of at least one object class ("prior probability *p(c)* of spring class *c*", pg 3); and

means and variances of object appearance maps ("means and variances of the sprite appearance maps", pg 3).

It would have been obvious to one of ordinary skill in the art at the time the invention was made for system of *Foote et al.* to include wherein the model parameters include: a prior probability of at least one object class; and means and variances of object appearance maps as taught by *Jojic et al.* to "focus on learning the appearances of multiple objects in multiple layers, over the entire video sequence.", *Jojic et al.*, pg 1 and to provide "probabilistic 2- dimensional appearance maps and masks of moving, occluding objects.", *Jojic et al.*, pg 1.

Regarding **claim 21**, while *Foote et al.* in view of *Jojic et al.* discloses the system of 20, *Foote et al.* in view of *Jojic et al.* do not teach wherein the model further comprises observation noise variances.

Jojic et al. teaches a learning flexible sprites in video layers wherein the model parameters include observation noise variances "the observation noise variances β", pg 3.

It would have been obvious to one of ordinary skill in the art at the time the invention was made for system of *Foote et al.* to include wherein the model further comprises observation noise variances as taught by *Jojic et al.* to "focus on learning the appearances of multiple objects in multiple layers, over the entire video sequence.", *Jojic et al.*, pg 1 and to provide "probabilistic 2- dimensional appearance maps and masks of moving, occluding objects.", *Jojic et al.*, pg 1.

Regarding **claims 25** and **26**, while *Foote et al.* discloses the computer-implemented process of claim 23, *Foote et al.* does not teach wherein the model parameters of each generative model includes

(i) an object class appearance map,

(ii) a prior probability of at least one object class, and

(iii) means and variances of that object class appearance map.

Jojic et al. teaches a learning flexible sprites in video layers wherein the model parameters includes (i) an object class appearance map, (ii) a prior probability of at least one object class, and (iii) means and variances of that object class appearance map (Section 5, "Interference and Learning", first paragraph, pg 3).

It would have been obvious to one of ordinary skill in the art at the time the invention was made for each generative model of *Foote et al.* to include (i) an object class appearance map, (ii) a prior probability of at least one object class, and (iii) means and variances of that object class appearance map as taught by *Jojic et al.* to "focus on learning the appearances of multiple objects in multiple layers, over the entire video sequence.", *Jojic et al.*, pg 1 and to provide "probabilistic 2- dimensional appearance maps and masks of moving, occluding objects.", *Jojic et al.*, pg 1.

<u>*Foote et al. in view of Eberman et al.*</u>

[15] **Claim 32** is rejected under 35 U.S.C. 103(a) as being unpatentable over *Foote et al.* in view of *Eberman et al.*

Regarding **claim 32**, claim 11 recites identical features as in claim 32. Thus, references/arguments equivalent to those presented above for claim 11 are equally applicable to claim 32.

Conclusion

[16] Any inquiry concerning this communication or earlier communications from the examiner should be directed to DAVID P. RASHID whose telephone number is (571)270-1578. The examiner can normally be reached Monday - Friday 7:30 - 17:00 ET.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Vikkram Bali can be reached on (571) 272-7415. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/David P. Rashid/
Examiner, Art Unit 2624

David P Rashid
Examiner
Art Unit 26244

/Vikkram Bali/
Supervisory Patent Examiner, Art Unit 2624